**NIA 프로젝트 주간보고서(9월 4주차)**

|  |  |  |  |
| --- | --- | --- | --- |
| **팀원** | 김경수, 이세림, 이지상 | | |
| **작성자** | 김경수, 이지상, 이세림 | **작성일** | 2021.09.29(수) |

1. **작성자**
2. **주요 진행사항**

|  |  |
| --- | --- |
| **이슈사항** | * **비지도학습을 고려한 추가 학습** * 좀 더 정확한 평가를 위해 E2ON 데이터셋과 AI-Hub 데이터셋을 섞어서 훈련 * **정확도 지표 조사** * acc\_score의 값이 combined model에 evaluate.py 에서의 y\_hat 출력값과 frame\_level\_mask의 값과 비교해서 퍼센트로 출력됨을 확인 |

1. **세부 진행사항**

|  |  |
| --- | --- |
| **요 일** | **업무내용** |
| **9월 24일(금)** | * **E2ON 데이터셋 평가 결과 분석(1)**   + acc값을 구하기 위해 평가 출력(실수)의 첫번째 자리를 반올림하는 numpy.round()을 사용하여 accuracy\_score()함수 사용      * + 반올림한 값을 출력하면 전부 0의 값이 나옴      * + 평가 출력값은 위처럼 실수값을 가지며 모두 0보다 작은 수, 반올림하면 0이 될 정도로 작은 수를 가짐   + 정수의 값을 가지고 acc를 구하는 accuracy\_score()함수가 아닌 다른 방법을 찾아봐야 할 필요가 있음 |
| **9월 25일(토)** | * **E2ON 데이터셋 평가 결과 분석(2)**      * + 일차적인 문제로 AUROC와 AUPR, Accuracy의 차이를 감소시키기 위한 의논 진행   + 기존에 만든 frame level masks는 전조현상을 포함하여 이상행동의 값을 1로 넣었으나 이를 수정해 이상행동 부분만 1의 값을 넣어 npy파일을 수정     <좌 AI-Hub dataset, 우 E2ON dataset>   * + 수정 후 AUROC와 AUPR, Accuracy의 차가 줄긴 했으나 두 값을 높일 필요가 있음 |
| **9월 26일(일)** | * **E2ON 데이터셋 평가 결과 분석(3)**   + MPED-RNN 모델은 비지도학습으로 정답 없이 데이터의 특징만 학습해 데이터간의 관계를 찾는 학습방법   + 학습 데이터를 비지도학습을 통해 프레임별 패턴을 학습하여 이상행동의 부분을 탐지하는 것으로 예상   + 정답이 없기 때문에 많은 학습 데이터가 필요하다고 판단해 E2ON 데이터만 가지고 학습해 평가하기에는 무리가 있다 판단해 AI-Hub 데이터셋과 E2ON 데이터셋을 한 번에 학습하여 E2ON 데이터셋 평가     <좌 E2ON 학습 후 평가, 우 AI-Hub+E2ON 학습 후 평가>   * + AUROC, AUPR, Accuracy 모두 거의 비슷한 값으로 출력 |
| **9월 27일(월)** | * **E2ON 데이터셋 평가 결과 분석(4)**   + 다양한 방법으로 해결하기 위해 frame level masks의 값을 변화를 주며 우선적으로 AUROC와 AUPR, Accuracy 값의 변화를 확인   **①**       * + 번호 순서대로 frame level masks의 값에 0을 하나씩 증가시켜 평가   + 0의 개수가 증가할수록 AUROC와 Accuracy는 증가하며, AUPR의 값은 점차 감소함을 볼 수 있음   + 이전에 언급했듯 round()함수가 y\_hat의 값을 모두 0으로 만드는 경우가 있어 frame level masks에 0이 증가할수록 그에 비례하여 Accuracy의 값이 증가   + 학습 데이터가 부족(①)하거나 기존 학습에 문제(②)가 있거나, 혹은 round함수를 이용하여 정수로 만들고 해당 값으로 Accuracy를 구하는 것이 아닌 다른 방법을 찾아보는(③) 등 논의 중 |
| **9월 28일(화)** | * **y\_hat 조사**      * y\_hat의 길이는 frame level mask의 프레임수와 똑같음 * 실제 y\_hat의 값은 0이 아니나 E2ON 데이터셋 일부의 경우 predict 결과 0.5 미만  의 값으로 이루어진 실수값이 나와 반올림시키면 전부 0의 값이 출력 * 이상행동의 비율이 많으면 학습시킬 때 0에 가까운 값이, 비율이 적으면 1에 가까  운 값이 나올거라 예상했지만 기존 학습 데이터 폴더에 있는 frame level masks 중  에 E2ON과는 다르게 이상행동의 비율이 높은 영상 존재 * 현재는 평가 결과가 나오는 과정에서 문제가 있다 보고 해당 문제를 고치려 노력중 |

1. **기타 보고사항**

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **차주  진행계획** | * 10월 1주차 진행계획  |  |  |  |  |  |  | | --- | --- | --- | --- | --- | --- | | **구분** | **10월 1일** | **10월 2일** | **10월 3일** | **10월 4일** | **10월 5일** | | **연구**  **내용** | 기존 학습 및 평가 과정 탐색 및 수정(1) | 기존 학습 및 평가 과정 탐색 및 수정(2) | 평가에 대한 accuracy 출력 방식 변경 | E2ON 데이터셋 추가 학습 및 평가 | 주간보고서 작성 및 발표 자료 준비 | | **산출물** |  | csv 또는 anomaly masks에 대한 오류 |  | 학습 및 평가 결과 | 주간 보고서및 발표자료 | |
| **특이사항** | * **E2ON 데이터셋 평가 출력 변환 문제**   + E2ON 데이터셋의 accuracy를 출력하기 위해 사용하는 accuracy\_score 함수는 정수값이 필요하며, 이를 위해선 일부 데이터셋을 이용한 모델 평가 결과를 정수로 바꾸는 수정이 필요   + MPED-RNN 모델에서 기존에 사용한 Avenue 데이터셋의 경우 일부 값이 1 이상이 존재했기 때문에 accuracy가 좋게 나왔지만 E2ON 데이터셋의 경우 대부분의 값이 0.1보다 작은 실수의 값으로 이루어져 있어 수정에 어려움 |